Doc. No. AM101
Appl. No. 09/894,018

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A61K — A2

(11) International Publication Number: **WO 96/22067**

(43) International Publication Date: 25 July 1996 (25.07.96)

(21) International Application Number: PCT/US95/16290

(22) International Filing Date: 15 December 1995 (15.12.95)

(30) Priority Data:
08/366,332 27 December 1994 (27.12.94) US

(71) Applicant: UNITED BIOMEDICAL, INC. [US/US]; 25 Davids Drive, Hauppauge, NY 11788 (US).

(72) Inventors: KUEBLER, Peter, J.; P.O. Box 2114, Ft. Lauderdale, FL 33303-2114 (US). NIXON, Douglas, F.; 112 Bayview Avenue, Northport, NY 11768 (US).

(74) Agent: LIN, Maria, C., H.; Morgan & Finnegan, L.L.P., 345 Park Avenue, New York, NY 10154-0053 (US).

(81) Designated States: AU, CA, FI, JP, NO, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: PEPTIDE RATCHET LIBRARIES FOR CTL-INDUCING VACCINES AND THERAPEUTICS

(57) Abstract

The present invention relates to ratchet libraries composed of related peptides synthesized simultaneously in a single peptide synthesis. Ratchet libraries are derived from a longer template peptide by sequentially "ratcheting" the template sequence into the shorter ratchet length and are used for cytotoxic T lymphocyte (CTL) induction or stimulation if the CTL epitope is known. If the CTL epitope is unknown, then the ratchet library can be used for identification of CTL epitopes. The ratchet libraries can be prepared from any protein sequence to which an immune CTL response is desired and can be formulated for delivery as a vaccine or therapeutic for the treatment or prevention of disease or malignancy. For example, a ratchet library can be used in the prevention and treatment of infectious or malignat diseases including HIV, influenza, malaria, breast, ovarian, lung and colon cancers.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia
AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CS	Czechoslovakia
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgystan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakhstan
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam

PEPTIDE RATCHET LIBRARIES FOR CTL-INDUCING VACCINES AND THERAPEUTICS

FIELD OF THE INVENTION

The present invention relates to ratchet libraries composed of related peptides synthesized simultaneously in a single peptide synthesis. Ratchet libraries are derived from a longer template peptide by sequentially "ratcheting" the template sequence into the shorter ratchet length and are used for cytotoxic T lymphocyte (CTL) induction or stimulation if the CTL epitope is known. If the CTL epitope is unknown, then the ratchet libary can be used for identification of CTL epitopes. The ratchet libraries can be prepared from any protein sequence to which an immune CTL response is desired and can be formulated for delivery as a vaccine or therapeutic for the treatment or prevention of disease or malignancy. For example, a ratchet library can be used in the prevention and treatment of infectious or malignant diseases including HIV, influenza, malaria, breast, ovarian, lung and colon cancers.

BACKGROUND OF THE INVENTION

The development of vaccines and therapeutics specifically designed to stimulate cytotoxic T lymphocytes (CTL) is needed. CTL are a vital component of the natural immune response against infectious organisms and malignant cells. CTL are $CD8^+$ thymus derived lymphocytes which appear early in an immune response and help in the elimination of, for example, virus-infected cells or tumor cells by lysis of the target cells and by secretion of chemical immunomodulators termed cytokines, such as interferons.

CTL have been detected following many viral infections, including HIV infection, and extensive evidence points to a major role for CTL in control of virus infections [McMichael et al. (1983) New Eng. J. Med.

301:13; Nixon et al. (1992) Immunology 76:515). For example, adoptive transfer of specific CTL to influenza [Taylor et al. (1986) Immunology 58:417] or paramyxovirus simian virus 5-infected mice, cleared the virus from the lungs [Young et al. (1990) J. Virol. 64:5403]. Tumor specific CTL have also been shown to clear tumors caused by mouse retroviruses [Cerundolo et al. (1987) Eur. J. Immunol. 17:173] and are also probably critical in the control of certain human malignancies. It has long been the aim of scientists to develop vaccines or therapeutics designed to specifically stimulate CTL immunity.

An essential step in the design of a CTL-inducing vaccine is in the identification of the antigenic sites to which CTL react. CTL recognize infected or malignant cells through the interaction of their specific T-cell receptor with a complex displayed on the surface of the target cell. The complex consists of an antigenic peptide specific to the virus or tumor, for example, and a major histocompatibility complex (MHC) class I molecule encoded by the Class I MHC genes of the host [Townsend et al. (1986) Cell 44:959]. Clusters of closely linked MHC alleles are characteristically inherited as a genetic unit termed the "haplotype". In general, individual MHC molecules associate with and present different antigenic protein fragments, so that one fragment of an antigenic protein is recognized by CTL of a specific MHC haplotype, while a different MHC haplotype requires another fragment of the antigen for recognition, i.e., recognition of individual antigenic fragments is MHC-restricted. As the MHC alleles are highly polymorphic between diverse genetic groups, a large number of distinct peptides may be needed to insure CTL stimulation across diverse human populations.

The exact fragment(s) of a virus or tumor antigen or other potential antigenic site (i.e., CTL epitope) recognized by a specific CTL was thought to be between 7

and 25 amino acids, but recent characterization of viral peptides naturally processed in virus-infected cells and displayed by Class I MHC molecules have identified the CTL epitopes as peptides of between 7 to 11 amino acids in length [Rötzschke et al. (1991) Immunology Today 12:447] with the majority of these peptides being of 9 amino acids (nonomers).

Identification of CTL epitopes in a protein sequence has been achieved by using synthetic peptides to map immunogenic sites. For example, several human CTL epitopes have been defined from HIV through an *in vitro* testing process of the human immune response to HIV infection [Nixon et al. (1988) Nature 336:484-487; Nixon et al. U.K. Patents GB 2,255,093, 2,273,709, 2,273,710]. While many HIV CTL epitopes have been identified in animals, few have been identifed in humans. However, because CTL epitopes are simultaneously recognized by a T-cell receptor that is specific for both the virally-encoded peptide and the host-encoded MHC for clearance of an infected cell to occur, CTL-epitopes are species specific. Hence, human CTL epitopes may not be reliably predicted from animal studies.

While vaccine development has led to successful vaccine against many infectious diseases, (e.g. polio, measles), there are several important pathogens for which vaccines are either ineffective or simply non-existant, for example HIV, hepatitis C virus (HCV) and herpes simplex virus (HSV). Moreover, there are no vaccines for treatment of malignancies.

The identification of CTL epitopes makes it feasible to design CTL-stimulating vaccines and other immunotherapeutics for prevention or treatment of disease by the clearance of virally-infected cells or malignant cancer cells. However, there remain at least four major problems associated with developing CTL-inducing vaccines and immunotherapeutics, namely, (1) the identification of

CTL epitopes or regions of proteins containing such epitopes, (2) the induction of specific CTL responses by peptides, (3) the need to accomodate MHC diversity with a large multiplicity of peptides and (4) the ability to provide for antigenic variation and escape mutations within the CTL epitopic regions.

With respect to the identification of CTL epitopes, progress has been made in identifying CTL epitopes from a number of target antigenic proteins,; however, there remain many proteins which contain potential CTL antigenic sites for which epitopes have not been identified. For example, the EBNA 1 protein of Epstein-Barr virus (EBV). The present invention provides a solution to this problem because the ratchet libraries can encompass extensive CTL antigenic regions and eliminate the need to precisely map CTL epitopes, or even to map the CTL epitopes at all. In addition, the ratchet libraries can be used to map antigenic sites.

Until recently, it was assumed that the induction of specific CTL responses by peptides could only be stimulated by endogenously-produced peptide fragments of endogenous proteins assembled into HLA Class I-antigenic peptide complexes on the cell surface. However, recent studies have demonstrated that CTL responses can be primed by administration of lipid-derivatized peptides [Deres et al. (1989) Nature 342:561], peptides in liposomes [Friede et al. (1994) Vaccine 12:791-797], or peptides admixed or conjugated to other biologically active substances [Shirai et al. (1994) J. Immunol. 152:549]. Hence, ratchet library peptides can be formulated into an appropriate vehicle to elicit CTL responses in vivo.

To accomodate genetic diversity and MHC restriction, the ratchet library peptides provide a major advance since several epitopes can be incorporated into the ratchet libraries rather than relying upon mixtures individually synthesized immunogenic peptides.

The ability to provide for antigenic variation and escape mutations within the CTL epitopic regions is another significant problem. Antigenic variation is a recurrent problem among certain pathogens contributing to unsuccessful or limited success of vaccines. Extensive antigenic variation, for example, is a hallmark of HIV (AIDS), rhinovirus (the common cold), influenza virus (flu), plasmodium falciparum (malaria). In addition, some tumors and infectious agents use utilized escape mutation to avoid immune surveillance. Ratchet libraries can be constructed to embody known antigenic variation and escape mutations to pre-empt these problems.

Hence, the ratchet library method of CTL induction provides a solution to obstacles in the development of vaccines and therapeutics for such pathogens or cancers.

SUMMARY OF THE INVENTION

This invention is directed to a ratchet library of peptides comprising at least one immunostimulatory cytotoxic T lymphocyte (CTL) epitope. The peptides are of length l. The sequences of the peptides in the library are determined from a template peptide of length from l+1 to n amino acids such that each position x in the library has all the amino acids present in the template peptide at positions x to n-(l-x), inclusive and the ratio of amino acids at each position x is determined by the relative prevalence of amino acids at that position x. In accordance herewith l is from about 7 to about 25 amino acids, preferably 8-10 and more preferably 9; n is from l+1 to about 100, preferably from l+1 to about 75 and more preferably from l+1 to about 50; and x is from 1 to l.

Moreover, if a position x is identified as part of an MHC-binding motif of a CTL epitope, then that position x in the ratchet library is fixed as one or more amino acids of the MHC-binding motif in an equimolar ratio.

The CTL epitopes of this invention are from a virus, bacterium, parasite, tumor antigen, allergen or other protein antigen. The ratchet library can be constructed from the template peptides of any one of SEQ ID NOS: 1-11. the peptides can have a covalently attached N-terminal tripalmitoyl-5-glycerol-cysteine moiety or be linked to a branched core sequence, polymerized or conjugated to a carrier molecule.

Another aspect of this invention provides a pharmaceutical or vaccine composition comprising the subject ratchet libraries including emulsion or a microparticle formulation with or without the addition of free Pam_3Cys or a derivative thereof. These compositions are useful in a method of treating or preventing a disease or a malignancy which comprises administering an amount of the ratchet library as a vaccine or pharmaceutical composition to a mammal effective to stimulate a CTL response against the disease or the malignancy associated with the CTL epitope present in the library.

Still another aspect of the invention is directed to a method of constructing a library of related peptides to provide a ratchet library which comprises identifying a template peptide; calculating a distribution of amino acids at each position x having those amino acids present in the template peptide at positions x to n-(l-x), inclusive, wherein l is from about 7 to about 25, n is from l+1 to about 100, and x is from 1 to l; and synthesizing said ratchet library.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 depicts malaria ratchet libraries 1 and 2 from *Plasmodium berghei* circumsporozoite (CS) protein and their construction. Fig 1A shows the template peptide with the known CTL epitope (CS 252-260) indicated by a box. Below the template peptide is the corresponding set of overlapping nonmer peptides. Fig. 1B provides an example

of a sequence alignment for malaria ratchet library 1 (top panel), its corresponding amino acid (AA) distribution by position (middle panel) and the percent of each amino acid at each position in the library (bottom panel).

Fig. 2 is a graphic illustration of malarial-specific CTL activity against CS 252-260. The graph shows the percent specific cell lysis as a function of the effector to target cell (E:T) ratio in mice immunized with 100 μg doses of malaria ratchet library 1 in microparticles.

Fig. 3 is a graphic illustration of malarial-specific CTL activity against CS 252-260. The graph shows the percent specific cell lysis as a function of E:T ratio in mice immunized with 1 mg doses of malaria ratchet library 2 in microparticles.

Fig. 4 is a graphic illustration of malarial-specific CTL activity against CS 247-266. The graph shows the percent specific cell lysis as a function of E:T ratio in mice immunized with 1 mg doses of malaria ratchet library 2 in microparticles.

Fig. 5 is a graphic illustration of malarial-specific CTL activity against CS 252-260. The graph shows the percent specific cell lysis as a function of E:T ratio in mice immunized with 10 μg doses of malaria ratchet library 1 as lipopeptides.

Fig. 6 is a graphic illustration of the lack of malarial-specific CTL activity against a self peptide. The graph shows the percent specific cell lysis as a function of E:T ratio in mice immunized with 100 μg doses of malaria ratchet library 1 as lipopeptides.

Fig. 7 depicts an MHC-restricted malaria ratchet library constructed from malarial ratchet library 1 (top panel). The middle panel shows the known anchor residues for four MHC haplotypes, K^d, D^b, K^b and L^d. These anchor residues are at positions 2 for K^d and L^d, 5 for D^b, K^b and L^d, 8 for K^b and 9 for K^d, D^b and L^d. The bottom panel

shows the resulting MHC-restricted malaria ratchet library.

Fig. 8 depicts an HIV ratchet library from a 35 amino acid sequence of the HIV-1 gp120 V3 loop region (residues 305-339). Fig. 8A provides the sequence of 15 HIV-1 variants from this region (top panel) and the corresponding SSAL from those sequences. The D^d restricted CTL epitope at gp120 amino acid positions 318-326 is indicated by a box. Fig. 8B provides the amino acid (AA) distribution by position (middle panel) and the percent of each amino acid at each position in the library (bottom panel) in a ratchet library constructed from the SSAL. Amino acids divergent from the consensus B sequence are shown as upper case letters and conserved amino acids in the consensus sequences are shown as lower case letters.

Fig. 9 depicts an HIV-1 gag peptide linear ratchet library containing a mouse HIV CTL epitope prepared from a 100 amino acid template (top panel). The amino acid (AA) distribution by position (middle panel) and the percent of each amino acid at each position in the library (bottom panel) is shown. The D^b restricted epitope (at gag residues 390-398) is indicated by the box.

Fig. 10 depicts an HIV-1 gag peptide linear ratchet library containing a mouse HIV CTL epitope prepared from a 40 amino acid template (top panel). The amino acid (AA) distribution by position (middle panel) and the percent of each amino acid at each position in the library (bottom panel) is shown. The D^b restricted epitope (at gag residues 390-398) is indicated by the box.

Fig. 11 depicts an HIV-1 gag peptide linear ratchet library containing a mouse HIV CTL epitope prepared from a 20 amino acid template (top panel). The amino acid (AA) distribution by position (middle panel) and the percent of each amino acid at each position in the library (bottom

panel) is shown. The D^b restricted epitope (at gag residues 390-398) is indicated by the box.

Fig. 12 is a graphic illustration of HIV-specific CTL activity against HIV-1 gag residues 390-398. The graph shows the percent specific cell lysis as a function of E:T ratio in mice immunized with 100 μg doses of the HIV ratchet library from the 40-mer template in an emulsion.

Fig. 13 depicts a mucin ratchet library from a 20-mer repeating sequence (top line). The next two lines illustrate two alternate template peptides for construction of this mucin ratchet library. The boxed residues are the additional sequences added at the termini to allow representation of all possible nonomers of the 20 amino acid repeat sequence. The amino acid (AA) distribution by position (middle panel) and the percent of each amino acid at each position in the library (bottom panel) is shown.

Fig. 14 depicts a mutant p53 ratchet library constructed from a template peptide of amino acids 124-151 (top panel). There are additional amino acids, which represent known p53 mutants incorporated at positions 9-12. The boxed residues are 10 amino acid CTL epitope identified in Balb/C mice. The amino acid (AA) distribution by position (middle panel) and the percent of each amino acid at each position in the library (bottom panel) is shown.

Fig. 15 depicts influenza ratchet library 1 constructed from a 25-mer template sequence of residues 139-163 of influenza A A/34/PR8 nucleoprotein (top panel). The known K^d-restricted epitope of residues 147-155 is indicated by a box. The amino acid (AA) distribution by position (middle panel) and the percent of each amino acid at each position in the library (bottom panel) is shown.

Fig. 16 depicts influenza ratchet library 2 from a template peptide which is a linkage of 3 CTL epitopes in order from N to C terminus of residues 50-58, 147-155 and

366-374 of the nucleoprotein. Each CTl epitope is indicated by a box. The amino acid (AA) distribution by position (middle panel) and the percent of each amino acid at each position in the library (bottom panel) is shown.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed to a ratchet library of peptides comprising at least one immunostimulatory cytotoxic T lymphocyte (CTL) epitope. The peptides of the ratchet library are of length l and the sequences of the peptides in the library are determined from a template peptide having a length from l+1 to n amino acids. Each position x in the library peptides has those amino acids which are present in the template peptide at positions x to n-(l-x), inclusive. Accordingly, the ratio of the individual amino acids at each position x is determined from the relative numbers (or prevalence) of the amino acids at that position x. In accordance with this invention, l is from about 7 to about 25, n is from l+1 to about 100, and x is from 1 to l.

For example, in a ratchet library of the above formula, position 1 contains all the amino acids of the template peptide at positions 1 to n-(x-1), position 2 contains all the amino acids of the template peptide at positions 2 to n-(x-2), position 3 contains all the amino acids of the template peptide at positions 3 to n-(x-3), ... , position x-1 contains all the amino acids of the template peptide at positions x-1 to n-1, and position x contains all the amino acids of the template peptide at positions x to n.

Fig. 1 provides an example of a malaria ratchet library, showing the relative ratio of amino acids at each position in the ratchet library derived from a longer malaria template peptide as well as the percentage of amino acids required for synthesis at each position of the ratchet library. More specifically, the malaria template

peptide is divided up into sequential 9-mers which are aligned at the amino terminus of the template peptide to create a set of overlapping peptides from which to calculate the ratchet library composition. After calculation of its composition, the ratchet library is prepared in a single synthesis based on the calculated amino acid distributions at each position.

The size of the ratchet l, or ratchet length, can be determined from the actual or approximate size of the target CTL epitope. CTL epitopes have been identified which vary in length from 7 to 25 residues. However, the majority of CTL peptides are from 8-10 amino acids, and many are 9 amino acids. While the actual size of the CTL inducing peptide is preferred to determine the length l of the ratchet library, e.g. 9 amino acids, the ratchet length can be determined by other means, including an arbitrary selection of size within the range of 7 to 25 amino acids.

The size of the template peptide ranges from l+1 to about 100 amino acids, and preferably from l+1 to about 75 amino acids, and more preferably from l+1 to about 50. Selection of the template peptide length is an important factor in determining the overall complexity of the ratchet library, so that shorter template peptides tend to yield less complex ratchet libraries, i.e., have fewer peptides in the library.

If the CTL epitope is known, then the template peptide should be of a length n such that the CTL epitope is flanked by sufficient adjoining sequences, preferably at least l-1, to insure that the CTL epitope is represented in the ratchet library. If the CTL epitope is not known, then the template peptide can have a length which covers a significant region of the protein being tested. Typically such a length can range from about 20 to about 100 residues, and preferably ranges up to 50 or 75 residues. The template peptide can also be selected on

the basis of clustering of epitopes, of hydrophobicity, of stretches containing basic amino acids or of another biological characteristic. Selection of a longer template peptide is useful in identifying unknown CTL epitopes.

MHC binding motifs have identified particular amino acids residues within CTL epitopes which are important in peptide binding to the MHC recepter. When the MHC binding motif is known for a particular CTL epitope, then the ratchet library can be simplified by replacing the calculated distribution of amino acids at a particular site with the ratio of known amino acids from the MHC binding motif at that site. An example of this is shown in Example 2. In another example, the proportion of amino acids within a ratchet library can be altered to reflect human HLA binding motifs. For example, human HLA binding motifs for 9-mer or 10-mer peptides typically have the designated amino acids at the indicated positions: for HLA-A2, leucine at position 2, valine or leucine at the C terminus; for HLA-B35, proline at postion 2, tyrosine at the C terminus; for HLA-B53, proline at position 2, phenylalanine or tryptophan at the C terminus; for HLA-B8, lysine at position 3, lysine at position 5, isoleucine at the C terminus; for HLA-B27, arginine at position 2, lysine or arginine at the C terminus; for HLA-B7, alanine at position 1, proline at position 2, arginine at position 3, leucine or valine at the C terminus; for HLA-A68, threonine or valine at position 2, arginine at the C terminus; for HLA-A3.1, isoleucine or leucine at position 2, phenylalanine at position 3, lysine or tyrosine at the C terminus; and for HLA-A11, isoleucine or leucine at position 2, lysine at the C terminus.

When the CTL epitope is contained in a multiple tandemly repeated sequence, then the template peptide length n can be equal to the length of the CTL epitope plus l-1 residues where the l-1 residues carboxy-terminal amino acids of the epitope are placed at its amino

terminus, or alternatively, the 1-1 amino-terminal residues of the epitope are placed at its carboxyl terminus, for example as shown in Fig. 5 for the mucin ratchet library.

To accomodate antigenic variation, the ratchet library can be constructed from a template peptide which is itself a "structured synthetic antigen library" or SSAL. SSALs are defined and exemplified in U.S. Serial No. 08/143,412, filed October 26, 1993, which is incorporated herein by reference. Briefly, the sequence of an SSAL is determined by aligning the primary amino acid sequences of a related family of CTL epitopes and identifying the invariant and variant loci within the alignment. The invariant loci generally represent the structural framework of the SSAL. The degeneracy within the SSAL is determined by the loci within the alignment that harbor different amino acid residue types relative to an arbitrary prototype sequence. After determining which amino acids are to be at each position, the degree of degeneracy for the multiresidue position in the SSAL library is determined from the number of variants each individual amino acid represents by one of three methods. Thus in a simple manner, the specific amino acids and their frequency of appearance at each position within the SSAL is defined by the primary sequences of the different CTL antigens or molecules in the alignment of multiple primary sequences.

The degeneracies for the variant amino acid positions used for an SSAL can be determined in one of three ways. In one method, the identity and ratio of residues is determined by the relative prevalence of the amino acids in a compilation of known sequences for the epitope. In another method the identity of the amino acids at the variant position is determined from the compilation of known sequences for the epitope but the ratio of amino acids is set to be equimolar. Finally, the identity and

ratio of amino acids at a variant position can be determined by a modification of the first method to provide a simplified SSAL. When some of the residue types are present at a low frequency (i.e., less than 5-10% representation), the complex SSALs are modified to ensure adequate representation of all variants. The process follows three rules: 1) any amino acid present in the primary sequence list at a proportion less than 10% is set at 5% to allow for adequate representation of all variable positions; 2) amino acids occurring at frequencies greater than or equal to 10% are rounded to the nearest 10% of prevalence. If the sum of percent prevalence exceeds 100%, the percent of the amino acid with the highest prevalence is correspondingly reduced so that the amino acids occuring at a given position are each represented in the SSAL but the total representation does not exceed 100%.

Once an SSAL is calculated, then it can be the "template peptide" for construction of a ratchet library in accordance with the formula such that each position x in the library peptides has those amino acids which are present in the template SSAL at positions x to n-(1-x), inclusive. In other words, each full composition and ratios of amino acids at each position x in the SSAL is "ratcheted" and used to caluculate the final distribution of amino acids in the ratchet library.

The ratchet libraries can be prepared as a ensemble of linear peptides. Similarly, it can be attached to a branched core sequences, conjugated to a carrier or polymerized.

These core sequences include dendritically branched cores, linear array type branched cores or randomly branched cores (e.g. poly-L-lysine). The branched cores can be composed of an amino acid or an amino acid analog having two amino groups and one carboxyl group, each group capable of forming a peptide bond linkage. Preferably

such amino acids are lysine or a lysine analog such as ornithine. The amino acid analog can be an α-amino acid, a ß-amino acid, or any other either natural or non-natural amino acid with two amino groups and one carboxyl group available for forming peptide bonds. Preferred branched peptides of the invention are dimers, tetramers and octamers, especially those having a branching core structure composed of lysine such as a heptalysine core. Similarly, the branched cores can contain other residues interspersed among the branching residues as depicted, for example, in Fig. 12 of U.S. Serial No. 143,412.

When branched ratchet libraries are made, the library can have a C-terminal methionine as the residue that is attached to the branched core. The methionine provides a cleavable site to facilitate analysis of the ratchet library.

In addition, the ratchet can have one or more lysine residues (added at the amino terminus) to increase peptide solubility, cysteine and haloacylated residues can be added to facilitate directed coupling to carrier molecules, and methionine can be added for cyanogen bromide cleavage if necessary. Pam_3Cys, or a similar lipid tail, can be added to create a lipopeptide.

The subject ratchet libraries can also be used to form conjugates, i.e., the ratchet library, either in branched or linear form, can be coupled directly or indirectly, by methods known in the art, to carriers such as bovine serum albumin (BSA), human serum albumin (HSA), or other proteins, red blood cells or latex particles. In another embodiment, a ratchet library can be polymerized to homo- or hetero-dimers or higher oligomers by cysteine oxidation, by induced disulfide cross-linking, or by use of homo- or hetero-functional multivalent cross-linking reagents.

As used herein, a CTL epitope is a fragment of an antigen which binds to the peptide-binding cleft of an MHC

molecule such that the fragment-MHC complex is recognized by a T cell antigen-specific receptor (TCR) and thereby stimulates a CTL response. CTL and T cell epitopes are reviewed in Encyclopedia of Immunology (Roitt et al., eds.), 1992, Academic Press Ltd, London, in Vol. I at pages 447-450 and pages 515-517, respectively.

The protein sequence selected for a ratchet library can be from a protein with known immunogenic CTL epitopes, or from a protein whose CTL-stimulating ability has not been determined, in which case the ratchet library method can be used to identify CTL epitopes. Ratchet libraries can be constructed from CTL epitopes (or putative CTL epitopes) of viruses, bacteria, parasites, tumor antigens, allergens, amino acid sequences deduced from an intron or exon/intron mixtures, or from abberrant proteins often associated with malignancy and generated by frameshift mutations (i.e. frameshift sequences), or any other proteins known to stimulate a CTL response. More specifically, ratchet libraries can be prepared from the following proteins or proteins from the listed organisms or diseases (with the cited references indicating known CTL response to those proteins): melanoma proteins [Bakker et al. (1994) J. Exp. Med. 179:1005] including MAGE-1, -2, and -3 [Gaugler et al. (1994) J.Exp. Med. 179:921]; proteins associated with renal cell carcinoma; proteins associated with colon carcinoma [Townsend et al. (1994) Nature: 371:662]; proteins associated with prostate cancer (malignant or benign) including PSA; tyrosinase [Brichard et al. (1993) J. Exp. med. 178:489]; oncogenes such as the HER-2/neu proto-oncogene; ras [Gedde-Dahl et al. (1994) Eur. J. Immunol. 24:410]; MUC1 [Barnd et al. (1989) Proc. Natl. Acad. Sci. USA 86:7159]; p53 [Mijman et al. (1994) Immunol. Lett. 40:171]; p16; TL [Morita et al. (1994) J. Exp. Med. 179:777]; proteins from HIV-1 or -2 including envelope, gag, pol, nef, tat, rev, vpx, vpu [Nixon et al. (1992)]; HTLV-I or -II including envelope,

gag, pol, pX and TAX [Jacobson et al. (1990) Nature 348:245; lymphocytic choriomeningitis virus of mice (LCMV) [Aebisher et al. (1991) Proc. Natl. Acad. Sci. USA 88:11047]; influenza A, B or C including PB1, PB2, PA, NS1, M1, NP, HA [McMichael et al. (1978) Eur. J. Immunol. 8:705]; Epstein-Barr virus (EBV) including TETA, EENL, EBNA3, EBNA1 and LMP [Brooks et al. (1993) J. Exp. Med. 178:879]; respiratory syncytia virus (RSV) [Bangham et al. (1985) J. Virol. 56:55]; hepatitis B virus (HBV) [Bertoletti et al. (1993) J. Virol. 677:2367]; hepatitis C virus (HCV) [Koziel et al. (1992) J. Immunol. 149:3339] ; herpes simplex virus (HSV) [Bonneau (1993) Virology 195:62]; cytomegalovirus (CMV) [Borysiewicz et al. (1988) Eur. J. Immunol. 18:269]; parainfluenza virus 1 including hemagglutinin, neuraminidase, phosphoprotein and nucleoprotein [Dave et al. (1994) Virology 199:376]; intracisternal A particle gag [de Bergeyck et al. (1994) Eur. J. Immunol. 24:2203]; bovine leukemia virus [Gatei et al. (1993) J. Virol. 67:1796]; papilloma viruses [Feltkamp et al. (1993 Eur. J. Immunol. 23:2242]; malaria including *P. falciparum, P. berghei, P. ovale, P. vivax, P. malaria* [Aggarwal et al. (1990) J. Exp. Med. 172:1083]; *Histoplasma capsulatum* [Deepe (1994) J. Immunol. 152:3491]; Listeria [Harty et al. (1992) J. Exp. Med. 175:1531]; Toxoplasmosis [Khan et al. (1994) J. Immunol. 152:1856]; *Trypanosoma cruzi*; Yersinia; *M. tuberculi*; *M. lepri*; *Pneumocystis carinii*; Kaposi's sarcoma or frameshift sequences [Townsend (1994)].

The preferred ratchet libraries of this invention are those libraries provided in the Examples and the Figures.

CTL responses can be measured by conventional techniques known to ordinarily skilled artisans, including, for example, the dye exclusion test and the Cr-release assay described in Encyclopedia of Immunology, supra at page 451. Another method to assay CTL is

described by McDonald et al. (1980) Immunol. Rev. 51:93-123.

The ratchet libraries are prepared by chemical synthesis using standard techniques well known in the art such as the solid-phase synthetic route pioneered by Merrifield. The coupling pf multiple amino acids at a given position is accomplished by providing a mixture of the desired amino acids at the ratios determined by the ratchet process. If necessary the ratio of amino acids in the mixture can be varied to account for different coupling efficiency of those amino acids.

Based on CTL induction of the ratchet libraries, they are useful in a vaccine composition to treat or prevent disease or malignancy in accordance with the source of the CTL epitope in the ratchet library. In other words an HIV ratchet library can be used as an HIV CTL vaccine (either as a vaccine component or as a therapeutic in the treatment of AIDS), an HCV ratchet library as an HCV CTL vaccine, an influenza ratchet library as a flu CTL vaccine, a mutant p53 ratchet library as a cancer CTL vaccine and the like.

For example, efforts to develop a malaria vaccine have been hampered by the complexity of the parasite life cycle and the inability for an antibody-inducing vaccine alone to provide sufficient efficacy. A CTL response to liver stage antigens of the malaria parasite *Plasmodium falciparum* has been recently reported [Hill et al. (1992) Nature 360:434]. Stimulation of the CTL response against the parasite appears necessary for an effective vaccine, since CTL can eliminate parasite-infected cells at the liver stage when the parasite load is low.

Vaccine compositions containing one or more distinct ratchet libraries can be introduced into normal subjects to stimulate production of CTL by immunization protocols known in the art. Similarly the subject ratchet libraries (one or more libraries) can be formulated in a vaccine

composition using adjuvants, pharmaceutically-acceptable carriers or other ingredients routinely provided in vaccine compositions. Adjuvants for use in this invention include incomplete Freunds' adjuvant (IFA), alum, lipidic amino acids and Pam_3Cys (see description in the Examples). These latter two adjuvants can be either covalently attached to the ratchet to produce a lipopeptide ratchet library or formulated together with the ratchet library for co-administration.

Vaccine formulations are readily determined by one of ordinary skill in the art and include formulations for immediate release and for sustained release. Formulations contemplated by this invention include microparticles, microcapsules, emulsions, liposomes, DMSO-glycerol and the like.

The present vaccines can be administered by any convenient route including subcutaneous, oral, intramuscular, intravenous, intra-dermal, intraocular, vaginal, trans-dermal or other parenteral or enteral route. Similarly the vaccines can be administered as a single dose or divided into multiple doses for administration.

The vaccine compositions of the instant invention contain an immunoeffective amount of a ratchet library to treat or prevent disease or malignancy associated with the source of the CTL epitope in that ratchet library. Preferred vaccine compositions are effective for CTL induction with respect to malaria, HIV, HCV, mucin, p53 and influenza and ther associated pathogenic conditions. Such compositions in dosage unit form can contain about 10 ng to about 2 mg of the peptide (or mixture of peptides) per kg body weight. When delivered in multiple doses, the dosage unit form is conveniently divided into the appropriate amounts per dosage.

Accordingly, another aspect of this invention provides a method of treating or preventing a disease or a

malignancy which comprises administering an amount of the library of Claim 1 to a mammal effective to stimulate a CTL response against the disease or malignancy associated with a CTL epitope present in said library. Based on the source of the ratchet library (i.e., the virus, bacterium or other organism from which the template peptide was derived or a protein associated with malignancy), then one skilled in the art can readily determine the amount needed for delivery to obtain the desired therapeutic result, that is, the amount of library to induce a CTL response of therapeutic benefit for the disease or condition under treatment. Typically these dosages ranges are as indicated above for the vaccine formulation. Likewise, one of ordinary skill in the art can readily determine an efficacious formulation for delivery of the ratchet library.

In another embodiment of this invention, a ratchet peptide can be used to identify CTL epitopes within a protein sequence. Hence, this invention is directed to a method of constructing a library of related peptides to provide a ratchet library which comprises identifying a template peptide; calculating a distribution of amino acids at each position x having those amino acids present in the template peptide at positions x to n-(l-x), inclusive, wherein l is from about 7 to about 25, n is from l+1 to about 100, and x is from 1 to l; synthesizing said ratchet library; and assaying said ratchet library for the ability to stimulate CTL activity. For example, the ratchet peptide is constructed and used to immunize animals, typically though not necessarily mice. The immunized animals are sacrificed and splenocytes removed and cultured in vitro with a pool of overlapping individual peptides that span the ratcheted template. The activate splenocytes are then tested on target cells pulsed with 50 μM (for example) pooled peptides, and if any CTL activity is present, the splenoytes are tested on

the individual peptides derived from the region. If a single peptide derived from the pool of overlapping peptides is recognized, a new CTL epitope has been identified.

The examples serve to illustrate the present invention and are not to be used to limit the scope of the invention.

EXAMPLE 1

Malaria Ratchet Libraries Generate Malaria-Specific CTL

A. General Methods

Ratchet libraries were synthesized by standard F-moc chemistry using solid phase peptide synthesis with an F-moc RINK MBHA resin [4-(2',4'-dimethoxyphenyl-Fmoc-aminomethyl)-phenoxyacetamido-norleucyl-MBHA resin; MBHA is methylbenzhydrylamine] according to manufacturer's instructions on an ABI Model 433 peptide synthesizer or similar model. The ratchet libraries were synthesized as linear peptides or as branched peptides using a heptalysyl core.

Ratchet libraries were formulated at the indicated concentrations and then used for immunization as microparticles, emulsions or lipopeptides.

Microparticles were prepared according to the water-in-oil-water solvent evaporation method described in U.S. Serial No. 08/263,841, filed June 22, 1994, which is incorporated herein by reference, using polylactide-co-glycolide polymer Resomer RG 505 (Boehringher Ingelheim). Microparticles containing 100 μg of ratchet library were suspended in 0.5 ml phosphate-buffered saline (PBS) for intraperitoneal immunization of mice on days 0, 10 and 20 followed by sacrifice of the animals 7 to 10 days later.

Emulsions were prepared so that the final preparations contained 100 μg of ratchet library and 50 μg Pam_3Cys-seryl--lysyl-lysyl-lysyl-lysyl (Pam_3Cys-SKKKK) (SEQ ID NO:12) in a volume of 0.5 mL unless indicated otherwise. To prepare the emulsion, 4 mg of ratchet library was dissolved in 16 mL H_2O and 240 mg of egg lecithin was dispersed therein by homogenization (Model STD 1 fitted with a 0.25" tubular head, Silverson Machines, East Longmeadow, MA) at 10,000 rpm for 5 min. Pam_3Cys-SKKK (2 mg) was mixed with 4 g soya oil and then added to the library mixture by homogenization at 10,000

rpm for 5 min. Further mixing was conducted by ultrasonic pulsation of the emulsion with an ultra sonic probe (Vibra cell, Sonics and Materialds, Inc., Danbury, CT). The emulsions (0.5 mL) were injected intraperitoneally into mice on days 0 and 10 followed by sacrifice of the animals 7 to 10 days later.

For lipopetide ratchet libraries, the peptides of the ratchet library were covalently coupled to Pam_3Cys (tripalmitoyl-5-glycerol-cysteine) as generally described (Deres et al.) to produce the corresponding lipopeptide ratchet library. The lipopeptide ratchet libraries (100 μg) were suspended in a 0.5 mL of 1% DMSO in glycerol and injected intraperitoneally into mice at day 0. The animals were sacrificed 7-9 days later.

Control mice were injected with 0.5 mL phosphate-buffered saline (PBS) using the corresponding injection schedule as that of the formualted ratchet library.

Upon sacrifice, the spleens were removed and splenocytes were pooled and cultured in vitro with the indicated peptide at a concentration of 1 μg/mL for one week to produce activated splenocytes.

CTL assays were then conducted according to the method of McDonald et al. (1980) as briefly described below.

B. Malarial Ratchet Libraries

A ratchet library was designed from a 20 amino acid sequence of residues 247-266 from the circumsporozoite protein of *Plasmodium berghei* (CS 247-266). This sequence contains the nonomer CTL epitope designated as CS 252-260 [Eberl et al. (1993) Int. Immunol. 5:1489-1492].

Malarial ratchet library 1 from the template peptide of SEQ ID NO:1 (Fig.1B, bottom panel) was prepared in linear form and formulated in microparticles at a final concentration of 200 μg/mL. Three BALB/C mice were immunized with 0.5 mL per injection as described above.

Pooled splenocytes were cultured with *Plasmodium berghei* CS peptide 252-260.

CS 252-260-specific CTL activity was then assayed. Briefly, H-2^d target cells (mouse mastocytoma cell line P815 or A20.1) were ^{51}Cr labeled for one h, washed and then incubated for one h in the presence of CS 252-260 peptide at 50 μM, in the presence of an unrelated control peptide, influenza nucleoprotein peptide NP 147-155 at 50 μM or in media (i.e., in the absence of a peptide antigen). CTL activity was determined by incubating these target cells with activated splenocytes (effector cells) for 4 hours in round-bottomed 96-well plates at a range of effector:target (E:T) ratios of 100:1, 50:1, 25:1 and 12.5:1 and measuring the release of ^{51}Cr. Cell lysis was calculated as per cent target cell lysis from the formula (E-M/T-M) x 100, where E = experimental ^{51}Cr release (cpm); M = ^{51}Cr release in presence of culture medium; and T = total ^{51}Cr released by 10% Triton X-100®.

The results are shown in Fig. 2 and indicate that significant CTL lysis was elicited in an MHC-restricted manner, since the K^d-restricted CS 252-260 peptide epitope sensitized target cells incubated with that peptide and not with any of the controls. Significant CTL lysis occurs if there is greater than 10% lysis above the control level of lysis at the highest E:T ratio. Splenocytes from control mice did not elicit specific CTL activity in any experiment.

Malaria ratchet library 2 (Fig. 1B, bottom panel) was synthesized as branched octameric peptides on a heptalysyl core and formulated in microparticles at a final concentration of 2 mg/mL and injected into mice as described above. Splenocytes were incubated and CTL activity was assayed as described above using E:T ratios of 100:1, 50:1 25:1 and 12.5:1. The results are shown in Fig. 3.

Malaria ratchet library 2 was formulated and injected into mice as described in the preceding paragraph to determine CTL activity against the 20 residue CS 247-266 peptide. Splenocytes were incubated with CS 252-260 peptide and CTL activity was determined as above, except that the target cells were incubated in the presence of 50 μM of CS 252-260 peptide, 50 μM of CS 247-266 peptide or in media using E:T ratios of 100:1, 50:1 25:1 and 12.5:1. The results are shown in Fig. 4.

Malaria ratchet library 1 was formulated and injected as a lipopeptide except that the lipopeptide ratchet library was formulated at a concentration of 20 μg/mL. Splenocytes were incubated CS 252-160 and CTL activity was assayed as described above using E:T ratios of 100:1, 50:1 25:1 and 12.5:1. The results are shown in Fig. 5.

Malaria ratchet library 1 was formulated and injected as a lipopeptide ratchet library at a concentration of 200 μg/mL. Splenocytes were incubated with CS 252-260 and CTL activity was as described above to determine CTL activity against a self peptide. For the CTL activity determination, the target cells were incubated in the presence of 50 μM of CS 252-260 peptide, 50 μM of a K^d-restricted self peptide (SYFPEITHI; SEQ ID NO:13) or in media using E:T ratios of 100:1, 50:1 25:1 and 12.5:1. The results are shown in Fig. 6.

These malaria ratchet libraries elicit malaria-specific CTL at immunogen doses ranging from 10 μg to 1 mg (Figs. 2-5). The CTL activity is MHC restricted and is elicited when presented with the processed epitope as a longer peptide (Fig. 4). CTL activity was absent when target cells were pulsed with a K^d restricted self-peptide to test for any auto-immune reaction. No significant lysis was seen on the targets pulsed with the self peptide.

EXAMPLE 2

MHC-Binding Motif Restricted Ratchet Library

Fig. 7, bottom panel, illustrates an MHC-restricted malaria ratchet library from the template peptide of SEQ ID NO:1. This library is constructed from malaria ratchet library 1 of Fig. 1B (also top panel of Fig. 7) by replacing those positions which contain anchor residues for the K^b, D^b, K^d, and L^d molecules (see below) with an equal proportion of the anchor residues at the position in question. The anchor residues are those amino acids which have been identified as necessary for binding to the MHC class 1 molecule for the given haplotype. The anchor residues for the indicated haplotypes are shown in the middle panel of Fig. 7 and the MHC-restricted malaria ratchet library is shown in the bottom panel of Fig. 7.

At each of these positions (i.e., positions 2, 5, 8 and 9), the ratchet incorporates only those anchor amino acids shown in the middle panel. Thus, position 2 contains 50% tyrosine Y and 50 % proline; position 5 contains 33% asparagine, 33% tyrosine and 33% phenylalanine; position 8 contains 100% leucine and position 9 contains 25% isoleucine, 25% leucine, 25% phenylalanine and 25% methionine. Thus, this ratchet has the anchor residues involved in binding MHC class I molecules and stimulating CTL.

EXAMPLE 3

Ratchet Libraries Can Accomodate Antigenic Diversity

An HIV ratchet library was constructed from the SSAL (Fig. 8A, bottom panel) of template peptide SEQ ID NO:2 and the additional 14 HIV-1 sequences (Fig. 8A) from a 35 amino acid sequence of the HIV-1 gp120 V3 loop region, the priniciple neutralizing domain known to have extensive sequence variability. This region contains a D^d restricted CTL epitope at amino acid positions 318-326. The antigenic diversity of this region is accomodated by taking 15 HIV-1 consensus sequences including the sequence HIV-MVP5188 and constructing an SSAL library where the identity and ratio of amino acids at each position is determined by the relative prevalence of amino acids in those 16 sequences. Next, it is the SSAL library which is "ratcheted" to yield the HIV-1 ratchet library shown in the bottom panel of Fig. 8.

EXAMPLE 4

Peptide Ratchet Libraries from Different Length Template Peptides

To examine the effect of template peptide length on the efficacy of CTL induction via ratchet libraries, three HIV-1 gag peptide linear ratchet libraries (Figs. 9-11, bottom panel) containing a mouse HIV CTL epitope were synthesized using template peptides of lengths 100, 40 or 20 amino acids of the gag sequence as shown in Figs. 9-11, respectively, and designated by SEQ ID NOS:3-5, respectively.

These libraries were formulated with 100 μg in 0.5 mL as microparticles, emulsions or lipopeptides and injected as described in Example 1 with the following modifications: The immunized mice were C57BL/6 mice. Activated splenocytes were prepared by culturing with 1 μg/mL HIV gag peptide 390-398 [Elvin et al. (1993) J. Immunol. Methods 158:161-171]. The target cells were EL4 and were incubated with 50 μM HIV gag peptide 390-398. The results are presented in Table 1. Fig. 12 show the specific cell lysis results for the emulsion formulation of the ratchet library from the 40 amino acid template.

Table 1

CTL Response from HIV-1 Ratchet Libraries

Formulation[a]	Pam_3Cys[b]	Template Peptide Length[c] 20	40	100
Lipopeptide	+[d]	+	±	++
Emulsion	+	±	+++	-
Microparticle	-	-	-	-

[a] The formulations are described in Example 1 and the experimental protocol in Example 4.

[b] Pam_3Cys is covalently bound to the ratchet library for the lipopeptide preparation but not in the emulsion preparation.

[c] The template peptide of the indicated amino acid length was ratcheted to 9-mers as described in Example 4.

[d] The symbols are as follows: +++, ++ and + mean significant specific target cell lysis in the indicated relative amounts with +++ as the most lysis; ±, inconclusive amount of specific cell lysis; -, no significant specific cell lysis.

EXAMPLE 5

Mucin Ratchet Libraries

Fig 13. (bottom panel) provides a mucin ratchet library constructed from the template peptide of SEQ ID NO:7 or SEQ ID NO:8. Mucin is a large, heavily glycosylated molecule expressed and secreted by ductal epithelial cells and tumors. Mucin consists of multiple copies of a 20 amino acid tandem repeat (SEQ ID NO:6) which appears to elicit non-MHC restricted CTL responses.

Because the 20-mer repeat will not contain every possible nonomer when used as a template, due to end effects in the ratchet, an alternative approach was used to generate all possible nonomers of the repeating 20-mer peptide. In this case the last eight carboxy-terminal amino acids of the 20-mer repeat peptide were placed at its amino terminus to yield a template peptide of 28 amino acids before calculation of the ratchet (Fig. 13, middle panel). Alternatively, the first eight amino-terminal amino acids were placed at its carboxyl terminus to yield a template peptide of 28 amino acids before calculation of the ratchet (Fig. 13, middle panel). With either method the calulated mucin ratchet library is the same.

EXAMPLE 6

Mutant p53 Ratchet Libraries

The protein p53 is a tumor supressor which fails to function effectively when mutated. More than 50% of human tumors contain cells which express a mutant form of p53, due to one or more point mutations in the protein. Class I mutations in p53 affect residues that directly contact DNA and include the residues lysine at postion 120, serine at position 241, arginine at position 248, arginine at position 273, alanine at position 276, cysteine at position 277 and arginine at positon 283. In this group, the mutations of arginine at positions 248 and 273 appear most frequently. Class II mutants affect residues that do not contact DNA but rather appear to have a role in stabilizing protein structure and include mutations of arginine at positions 175 and 249.

This example provides four peptide ratchet libraries containing four hot spots of mutation of p53.

The four ratchets are designed around mutation hot spots in the protein: (1) template peptide of amino acids 124-151 (SEQ ID NO:9) as a 10-mer ratchet library (Fig. 14, bottom panel); (2) template peptide of amino acids 166-187 as a 9-mer ratchet library; (3) template peptide of amino acids 228-256 as a 9-mer ratchet library; and (4) template peptide of amino acids 264-289 as a 9-mer ratchet library.

The ratchet library from the 166-187 template is especially useful for treating colon cancer since this mutation is frquently encountered with this malignancy.

EXAMPLE 7

Influenza Ratchet Libraries

Existing influenza vaccines are complex to design and manufacture as the prevalent strain of influenza can rapidly change and a vaccine designed to stimulate antibodies against influenza A of strain one may not be effective in eliciting cross-reactive immunity against strain two. The CTL response has been shown to be effective against influenza in animal models and in humans, and the addition of an influenza specific CTL component to existing vaccines, or a CTL inducing vaccine alone, would dramatically broaden protection against many strains of influenza.

To provide a vaccine capable of stimulating a CTL response against influenza, known CTL epitopes can be ratcheted. Fig. 15 (bottom panel) shows influenza ratchet library 1 from a 25-mer template peptide of residues 139-163 (SEQ ID NO:10) of influenza A A/34/PR8 nucleoprotein. This library encompasses the known K^d-restricted epitope of residues 147-155. Fig. 16 (bottom panel) shows influenza ratchet library 2 constructed from the template peptide of SEQ ID NO:11 which is a linkage of 3 CTL epitopes in order from N to C terminus of residues 50-58, 147-155 and 366-374 of the nucleoprotein.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Kuebler, Peter J.
Nixon, Douglas F.

(ii) TITLE OF INVENTION: Peptide Ratchets for Vaccines and Therapeutics

(iii) NUMBER OF SEQUENCES: 13

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: M. Lisa Wilson
(B) STREET: 25 Davids Drive
(C) CITY: Hauppauge
(D) STATE: NY
(E) COUNTRY: USA
(F) ZIP: 11788

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Wilson, M. Lisa
(B) REGISTRATION NUMBER: 34,045
(C) REFERENCE/DOCKET NUMBER: 2012

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (516)273-2828
(B) TELEFAX: (516)273-1717

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Asn Asn Asn Asp Asp Ser Tyr Ile Pro Ser Ala Glu
 1               5                  10
Lys Ile Leu Glu Phe Val Lys Gln
        15                  20
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile
 1               5                  10
His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr Gly
        15                  20
Glu Ile Ile Gly Asp Ile Arg Gln Ala His Cys
 25                  30                  35
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 100 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Ile Arg Gln Gly Pro Lys Glu Pro Phe Arg Asp Tyr
 1               5                  10
Val Asp Arg Phe Tyr Lys Thr Leu Arg Ala Glu Gln
        15                  20
Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr
 25                  30                  35
Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr
            40                  45
Ile Leu Lys Ala Leu Gly Pro Ala Ala Thr Leu Glu
     50                  55                  60
Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly Pro
                65                  70
Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
        75                  80
Gln Val Thr Asn Ser Ala Thr Ile Met Met Gln Arg
 85                  90                  95
```

Gly Asn Phe Leu
100

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly Pro
1 5 10
Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
15 20
Gln Val Thr Asn Ser Ala Thr Ile Met Met Gln Arg
25 30 35
Gly Asn Phe Leu
40

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Glu Ala Met Ser Gln Val Thr Asn Ser Ala Thr Ile
1 5 10
Met Met Gln Arg Gly Asn Phe Leu
15 20

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Pro Asp Thr Arg Pro Ala Pro Gly Ser Thr Ala Pro
1 5 10

Pro Ala His Gly Val Thr Ser Ala
15 20

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Pro Ala His Gly Val Thr Ser Ala Pro Asp Thr Arg
1 5 10
Pro Ala Pro Gly Ser Thr Ala Pro Pro Ala His Gly
15 20
Val Thr Ser Ala
25

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Pro Asp Thr Arg Pro Ala Pro Gly Ser Thr Ala Pro
1 5 10
Pro Ala His Gly Val Thr Ser Ala Pro Asp Thr Arg
15 20
Pro Ala Pro Gly
25

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Cys Thr Tyr Ser Pro Ala Leu Asn Lys Met Phe Cys
1 5 10

Gln Leu Ala Lys Thr Cys Pro Val Gln Leu Trp Val
15 20
Asp Ser Thr Pro
25

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Trp His Ser Asn Leu Asn Asp Ala Thr Tyr Gln Arg
1 5 10
Thr Arg Ala Leu Val Arg Thr Gly Met Asp Pro Arg
15 20
Met
25

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Ser Asp Tyr Glu Gly Arg Leu Ile Thr Tyr Gln Arg
1 5 10
Thr Arg Ala Leu Val Ala Ser Asn Glu Asn Met Glu
15 20
Thr Met
25

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1
(D) OTHER INFORMATION: /note= "N-terminal tripalmitoyl-5-glycerol"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Cys Ser Lys Lys Lys Lys
1 5

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Ser Tyr Phe Pro Glu Ile Thr His Ile
1 5

We Claim:

1. A ratchet library of peptides comprising at least one immunostimulatory cytotoxic T lymphocyte (CTL) epitope wherein said peptides are of length l; the sequences of said peptides in said library being determined from a template peptide of length from l+1 to n amino acids such that each position x in the library has all the amino acids present in said template peptide at positions x to n-(l-x), inclusive; the ratio of amino acids at each position x being determined by the relative prevalence of amino acids at that position x; and wherein l is from about 7 to about 25 amino acids, n is from l+1 to about 100, and x is from 1 to l.

2. The library of Claim 1 wherein n is from l+1 to about 75.

3. The library of Claim 1 wherein n is from l+1 to about 50.

4. The library of Claim 1 wherein l is from 8 to 10.

5. The library of Claim 1 wherein l is 9.

6. The library of Claim 1 wherein if a position x is identified as part of an MHC-binding motif of a CTL epitope, then that position x is fixed as one or more amino acids of said MHC-binding motif in an equimolar ratio.

7. The library of any one of Claims 1 to 6 wherein said CTL epitope is from a virus, bacterium, parasite, tumor antigen, allergen or other protein antigen.

8. The library of Claim 1 to 6 wherein said CTL epitope is from a melanoma protein including MAGE-1, -2, and -3; a renal cell carcinoma protein; a colon carcinoma protein; a prostate cancer protein (malignant or benign) including PSA; tyrosinase; an oncogene such as HER-2/neu proto-oncogene; ras; MUC1; p53; p16; TL; an HIV-1 or HIV-2 protein including envelope, gag, pol, nef, tat, rev, vpx or vpu; an HTLV I or II protein including envelope, gag, pol, pX or TAX; lymphocytic choriomeningitis virus of

mice; influenza A, B or C including PB1, PB2, PA, NS1, M1, NP or HA; an Epstein-Barr virus protein including TETA, EENL, EBNA3, EBNA1 or LMP; respiratory syncytia virus; hepatitis B virus; hepatitis C virus; herpes simplex virus; cytomegalovirus; a parainfluenza virus 1 protein including hemagglutinin, neuraminidase, phosphoprotein or nucleoprotein; intracisternal A particle gag; bovine leukemia virus; papilloma viruses; a malaria protein including proteins from *P. falciparum, P. berghei, P. ovale, P. vivax, P. malaria; Histoplasma capsulatum;* Listeria; Toxoplasmosis; *Trypanosoma cruzi;* Yersinia; *M. tuberculi; M. lepri; Pneumocystis carinii;* Kaposi's sarcoma or frameshift sequences.

9. The library of Claim 1 wherein said template peptide is any one of SEQ ID NOS: 1-11.

10. The library of any one of Claims 1-6 or 9 wherein said peptides of said library have a covalently attached N-terminal tripalmitoyl-5-glycerol-cysteine moiety.

11. The library of any one of Claims 1-6 or 9 wherein said peptides are linked to a branched core sequence, are polymerized or are conjugated to a carrier molecule.

12. A pharmaceutical or vaccine composition comprising the library of any one of Claims 1-6 or 9, or a pharmaceutically acceptable salt thereof, and a pharmaceutically acceptable carrier.

13. The composition of Claim 12 wherein said composition is an emulsion or a microparticle formulation.

14. The composition of Claim 12 wherein said formulation also comprises tripalmitoyl-5-glycerolcyteine or a derivative thereof.

15. A pharmaceutical or vaccine composition comprising the library of Claim 7, or a pharmaceutically acceptable salt thereof, and a pharmaceutically acceptable carrier.

16. The composition of Claim 15 wherein said composition is an emulsion or a microparticle formulation.

17. The composition of Claim 15 wherein said formulation also comprises tripalmitoyl-5-glycerolcyteine or a derivative thereof.

18. A method of treating or preventing a disease or a malignancy which comprises administering an amount of said composition of Claim 12 to a mammal effective to stimulate a CTL response against said disease or said malignancy associated with said CTL epitope present in said library.

19. A method of treating or preventing a disease or a malignancy which comprises administering an amount of said composition of Claim 15 to a mammal effective to stimulate a CTL response against said disease or said malignancy associated with said CTL epitope present in said library.

20. A method of constructing a library of related peptides to provide a ratchet library which comprises identifying a template peptide; calculating a distribution of amino acids at each position x having those amino acids present in the template peptide at positions x to n-(l-x), inclusive, wherein l is from about 7 to about 25, n is from l+1 to about 100, and x is from 1 to l; and synthesizing said ratchet library.

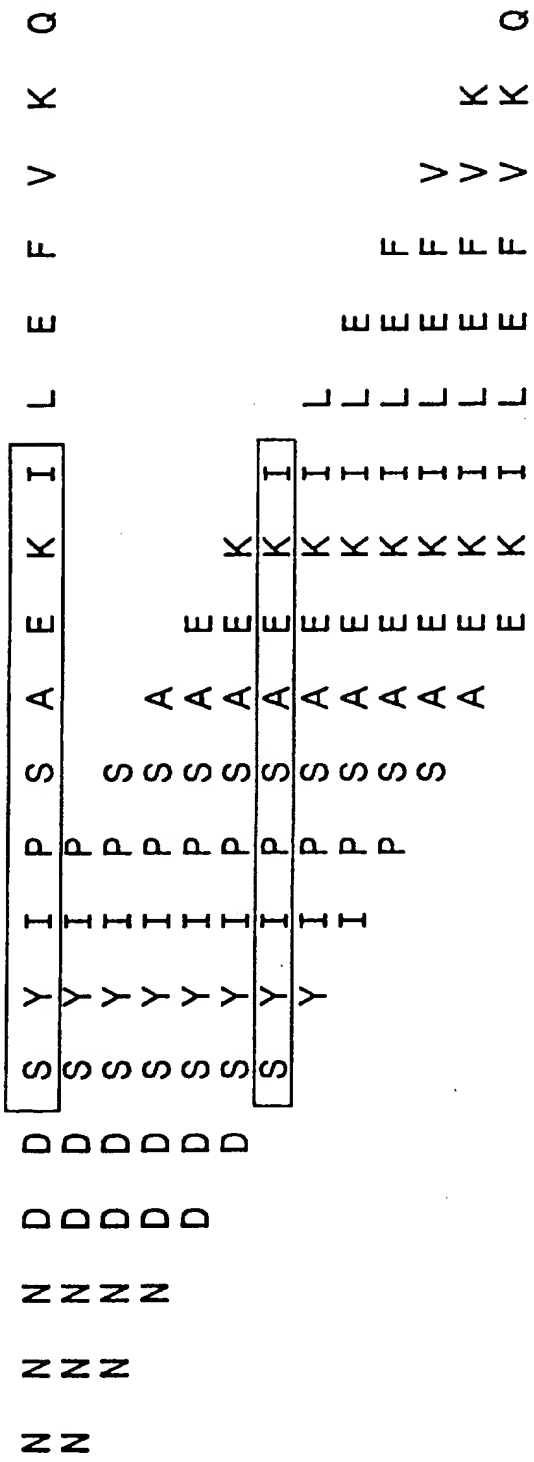


Figure 1A

Sequence Alignment by Position

1	2	3	4	5	6	7	8	9
N	N	N	D	D	S	Y	I	P
N	N	D	D	S	Y	I	P	S
N	D	D	S	Y	I	P	S	A
D	D	S	Y	I	P	S	A	E
D	S	Y	I	P	S	A	E	K
S	Y	I	P	S	A	E	K	I
Y	I	P	S	A	E	K	I	L
I	P	S	A	E	K	I	L	E
P	S	A	E	K	I	L	E	F
S	A	E	K	I	L	E	F	V
A	E	K	I	L	E	F	V	K
E	K	I	L	E	F	V	K	Q

AA Distribution by Position

AA residue	1	2	3	4	5	6	7	8	9
N	3	2	1						
D	2	2	2	2	1				
S	2	2	2	2	2	2	1	1	1
Y	1	1	1	1	1	1	1		
I	1	1	2	2	2	2	2	2	1
P	1	1	1	1	1	1	1	1	1
A	1	1	1	1	1	1	1	1	1
E	1	1	1	1	2	2	2	2	2
K		1	1	1	1	1	1	2	2
L				1	1	1	1	1	1
F						1	1	1	1
V							1	1	1
Q									1

AA Percentage by Position

AA residue	1	2	3	4	5	6	7	8	9
N	25	17	8						
D	17	17	17	17	8				
S	17	17	17	17	17	17	8	8	8
Y	8	8	8	8	8	8	8		
I	8	8	17	17	17	17	17	17	8
P	8	8	8	8	8	8	8	8	8
A	8	8	8	8	8	8	8	8	8
E	8	8	8	8	17	17	17	17	17
K		8	8	8	8	8	8	17	17
L				8	8	8	8	8	8
F						8	8	8	8
V							8	8	8
Q									8

Figure 1B

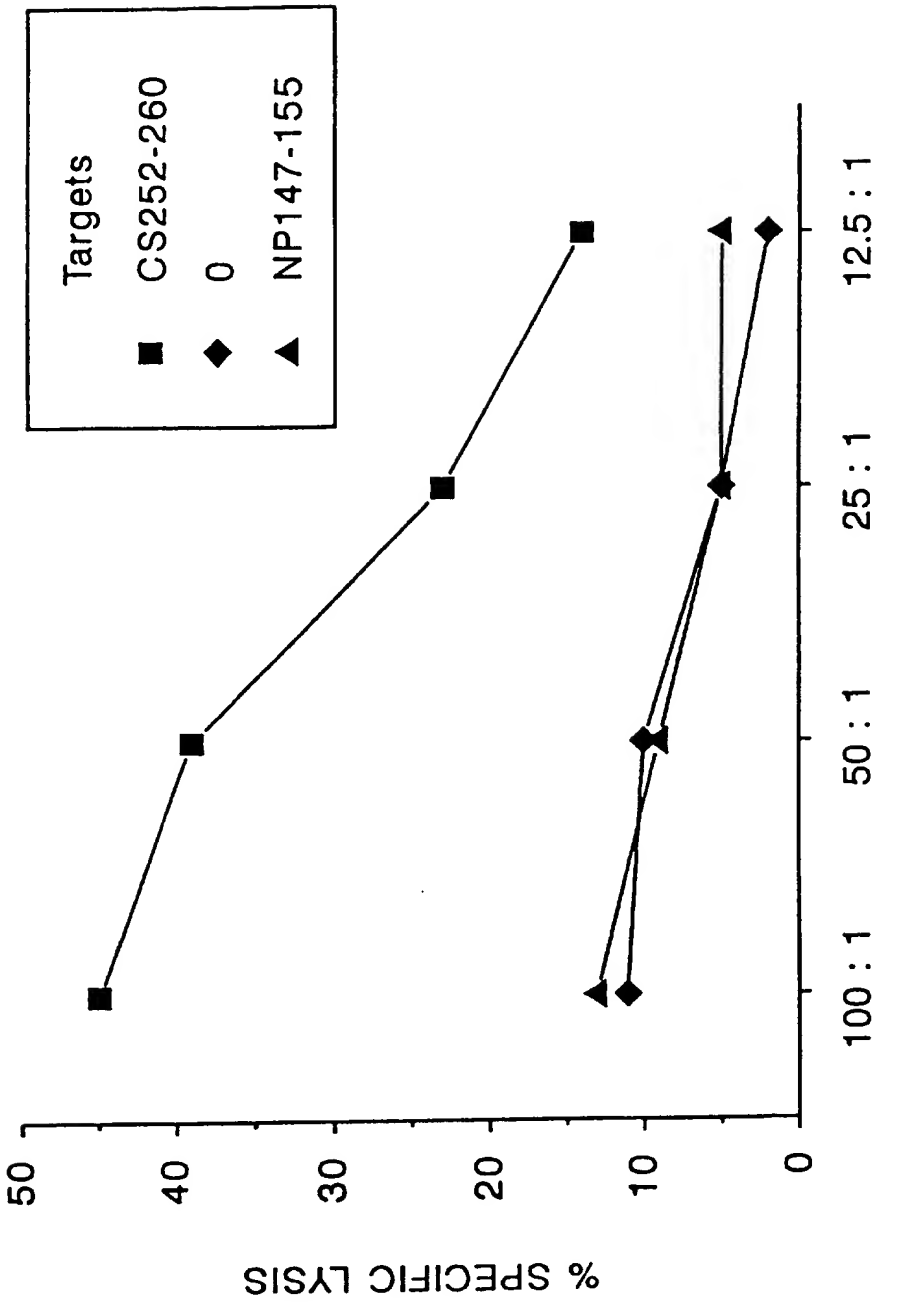

Targets
CS252-260
0
NP147-155
% SPECIFIC LYSIS
50
40
30
20
10
0
100 : 1
50 : 1
25 : 1
12.5 : 1
E:T RATIO

Figure 2

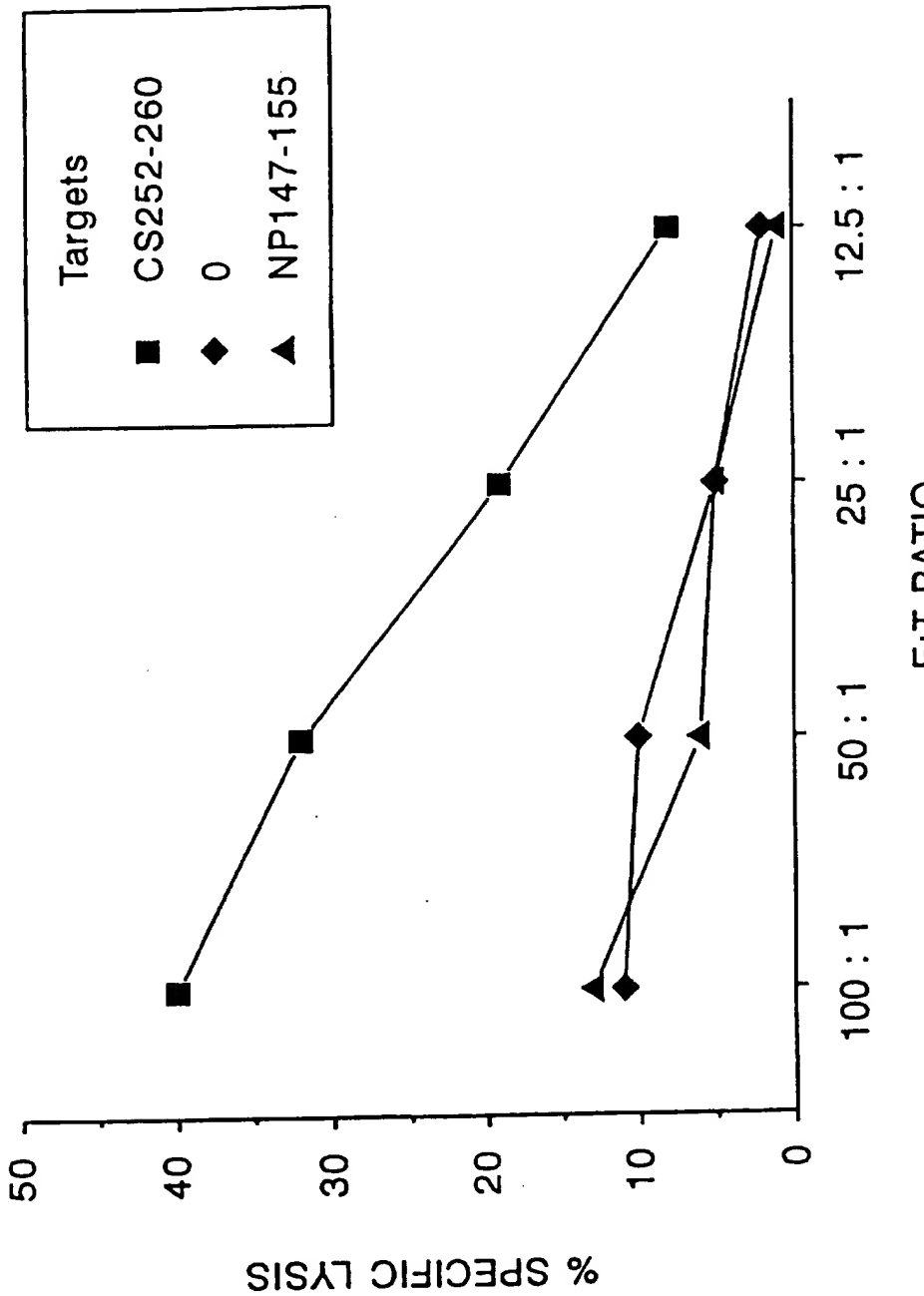

Targets
CS252-260
0
NP147-155
% SPECIFIC LYSIS
50
40
30
20
10
0
100 : 1
50 : 1
25 : 1
12.5 : 1
E:T RATIO

Figure 3

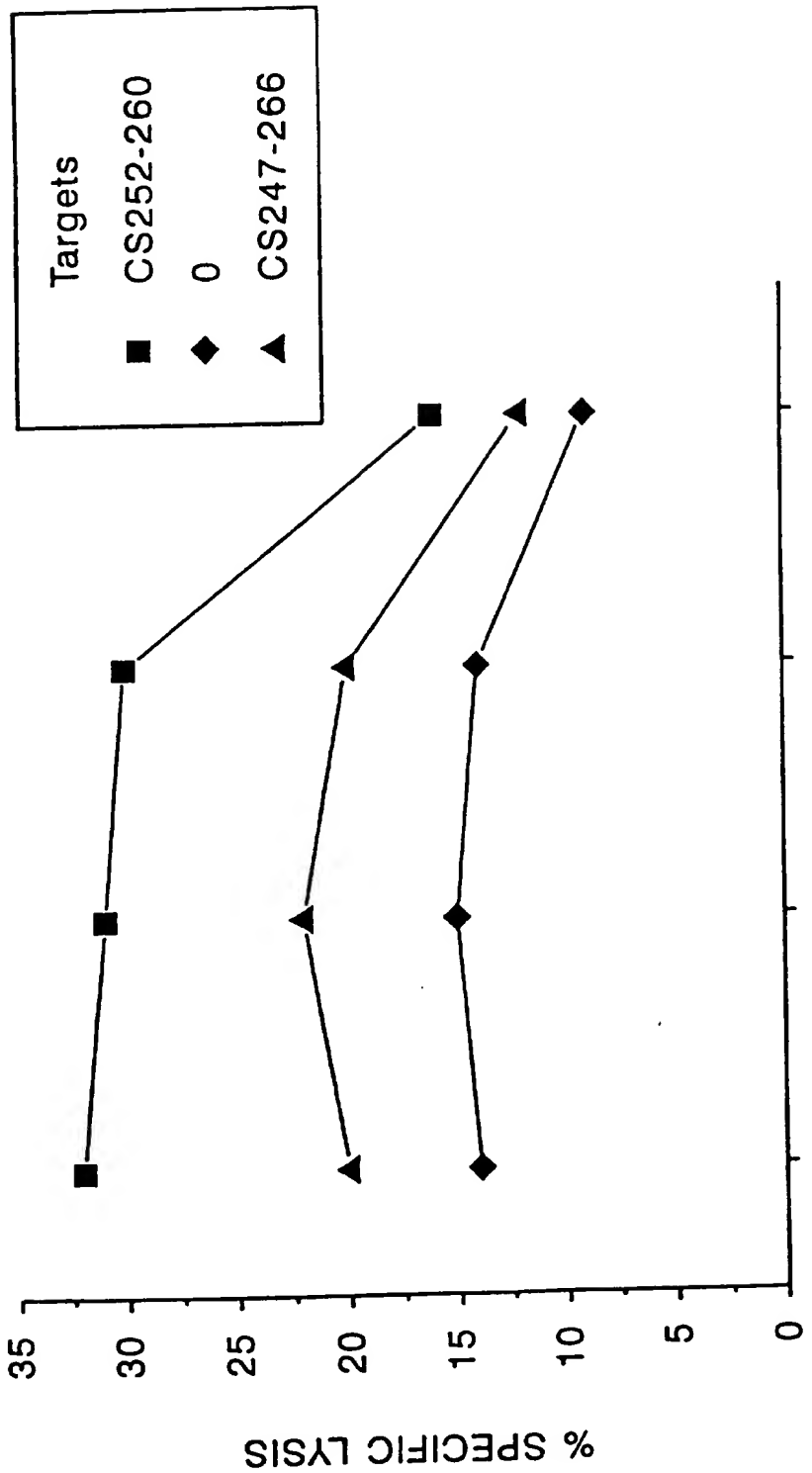

Targets
CS252-260
0
CS247-266
% SPECIFIC LYSIS
35
30
25
20
15
10
5
0
100 : 1
50 : 1
25 : 1
12.5 : 1
E:T RATIO

Figure 4

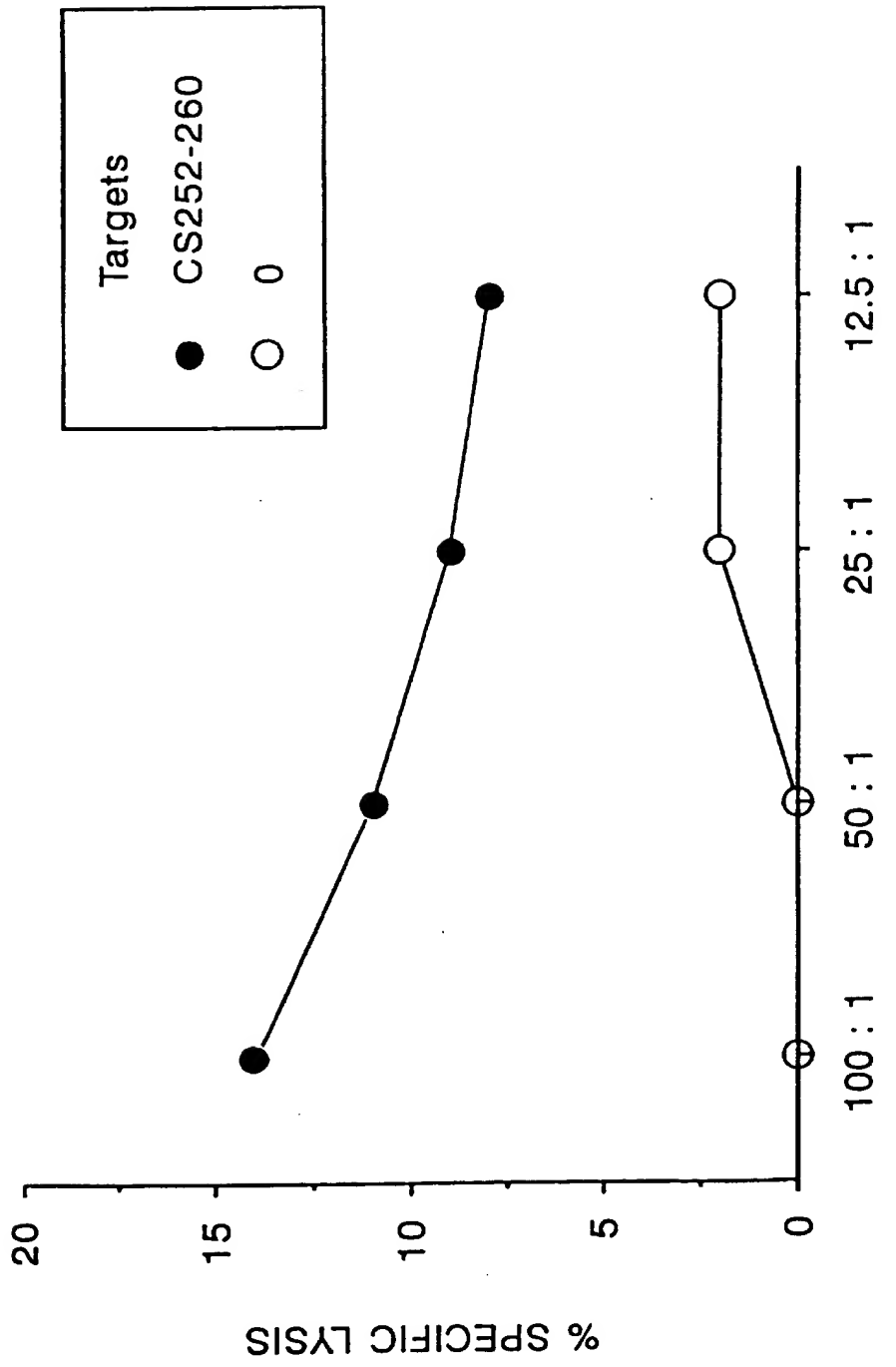

Targets
CS252-260
0
% SPECIFIC LYSIS
20
15
10
5
0
100 : 1
50 : 1
25 : 1
12.5 : 1
E:T RATIO

Figure 5

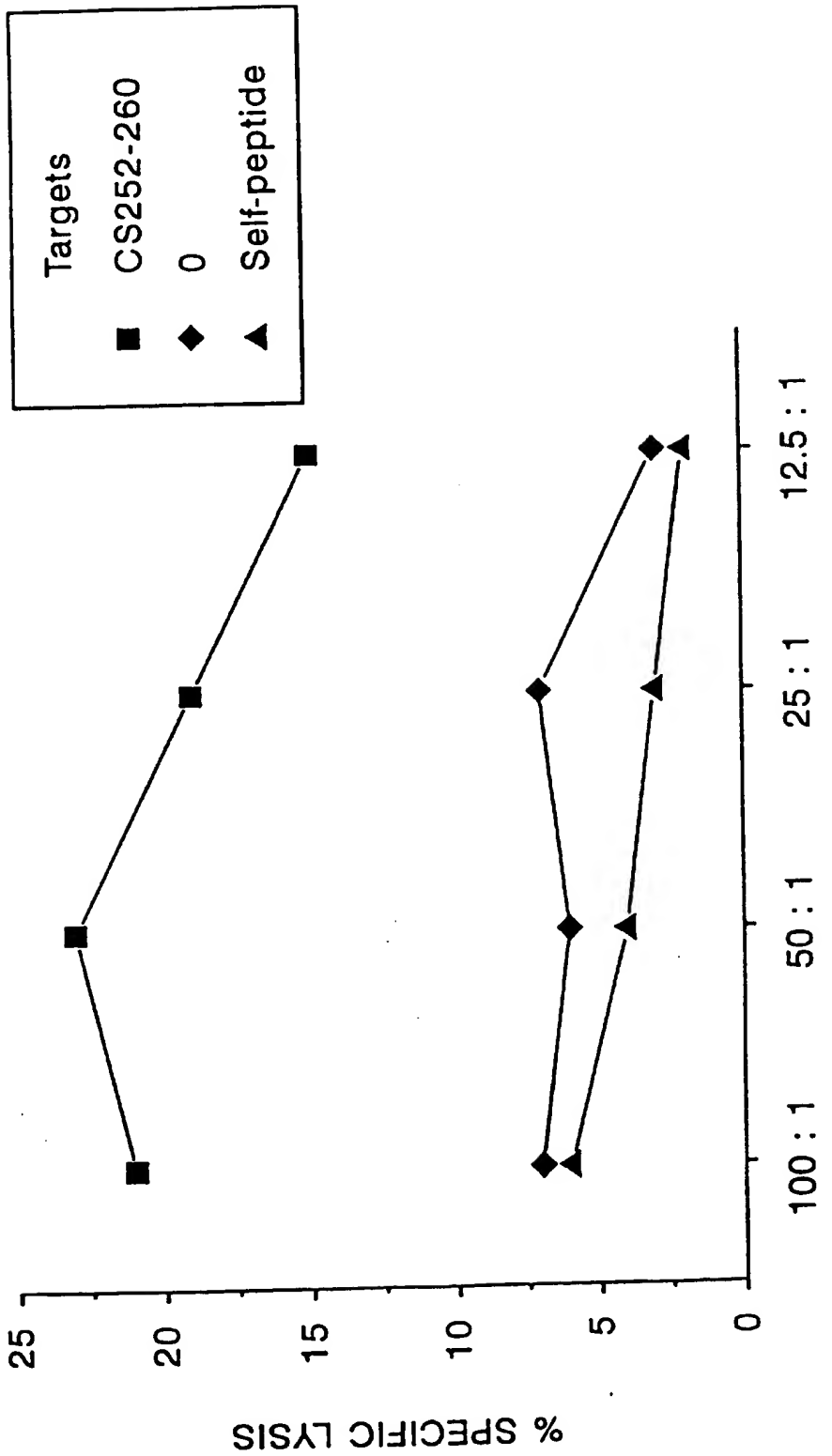

Targets
CS252-260
0
Self-peptide
% SPECIFIC LYSIS
25
20
15
10
5
0
100 : 1
50 : 1
25 : 1
12.5 : 1
E:T RATIO

Figure 6

AA Percentage by Position

AA residue	1	2	3	4	5	6	7	8	9
N	25	17	8						
D	17	17	17	17	8				
S	17	17	17	17	17	17	8	8	8
Y	8	8	8	8	8	8	8		
I	8	8	17	17	17	17	17	17	8
P	8	8	8	8	8	8	8	8	8
A	8	8	8	8	8	8	8	8	8
E	8	8	8	8	17	17	17	17	17
K		8	8	8	8	8	8	17	17
L				8	8	8	8	8	8
F						8	8	8	8
V							8	8	8
Q									8

AA Restrictions by Position

Haplotype									
K^d		Y							I L
D^b					N				M
K^b					F Y			L	
L^d		P							M L F

Restricted AA Percentage by Position

AA residue	1	2	3	4	5	6	7	8	9
N	25		8		33				
D	17		17	17					
S	17		17	17		17	8		
Y	8	50	8	8	33	8	8		
I	8		17	17		17	17		25
P	8	50	8	8		8	8		
A	8		8	8		8	8		
E	8		8	8		17	17		
K			8	8		8	8		
L				8		8	8	100	25
F					33	8	8		25
V							8		
Q									
M									25

Figure 7

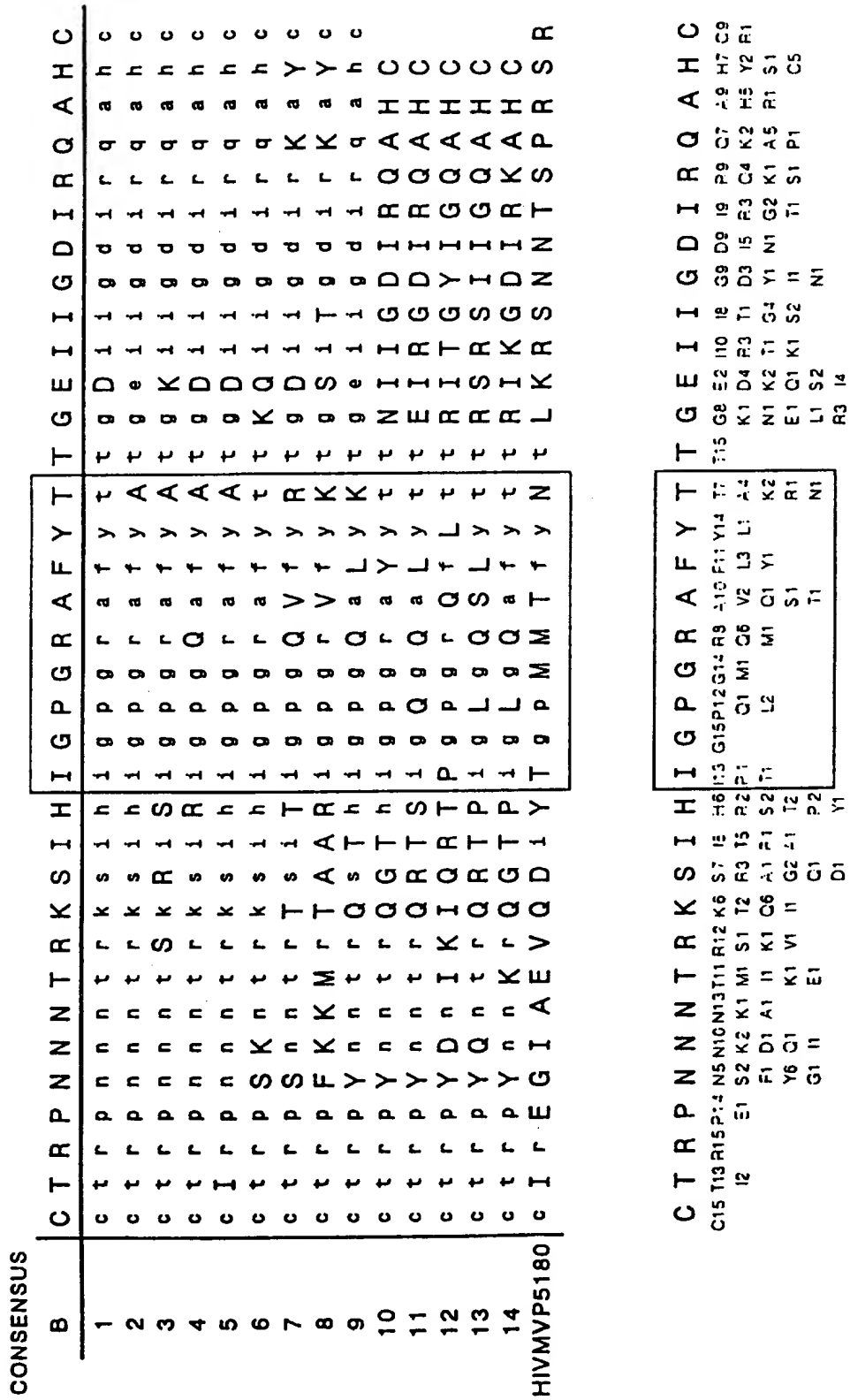

CONSENSUS
B
C T R P N N T R K S I H I G P G R A F Y T T G E I I G D I R Q A H C
1
2
3
4
5
6
7
8
9
10
11
12
13
14
HIVMVP5180

Figure 8A

AA Distribution by Position

AMINO ACID	1	2	3	4	5	6	7	8	9
C	15							5	9
T	59	59	46	47	47	47	47	47	36
R	48	48	48	36	45	45	46	46	45
P	29	29	29	29	16	16	16	16	16
I	48	49	42	51	51	51	50	50	49
K	17	17	17	17	18	18	16	15	14
E	5	5	5	5	4	4	4	4	3
N	30	31	30	30	30	25	15	2	2
F	12	12	12	12	12	11	11	11	11
Y	22	23	23	23	23	17	17	17	17
G	44	53	53	55	55	54	54	54	54
H	6	6	6	6	6	6	11	11	11
L	7	7	7	7	7	7	7	7	7
Q	17	17	17	17	21	21	20	20	20
S	17	17	17	17	18	16	16	16	16
D	6	9	18	18	18	18	17	17	17
A	17	17	17	17	17	17	26	25	25
V	3	3	3	3	3	3	3	3	3
M	3	3	3	3	3	3	3	3	2

AA Percentage by Position

AMINO ACID	1	2	3	4	5	6	7	8	9
C	4							1	3
T	15	15	12	12	12	13	13	13	10
R	12	12	12	9	12	12	12	13	13
P	7	7	7	7	4	4	4	4	5
I	12	12	11	13	13	13	13	13	14
K	4	4	4	4	4	4	4	4	4
E	1	1	1	1	1	1	1	1	1
N	7	7	8	8	8	7	4	1	1
F	3	3	3	3	3	3	3	3	3
Y	5	5	6	6	6	5	5	5	5
G	11	13	13	14	14	14	14	15	15
H	2	2	2	2	2	2	3	3	3
L	2	2	2	2	2	2	2	2	2
Q	4	4	4	4	5	5	5	5	6
S	4	4	4	4	4	4	4	4	5
D	2	2	5	5	5	5	5	5	5
A	5	5	5	5	5	5	7	7	7
V	1	1	1	1	1	1	1	1	1
M	1	1	1	1	1	1	1	1	1

Figure 8B

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
I	R	Q	G	P	K	E	P	F	R	D	Y	V	D	R	F	Y	K	T	L	R	A	E	Q	A

26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
S	Q	E	V	K	N	W	M	T	E	T	L	L	V	Q	N	A	N	P	D	C	K	T	I	L

51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
K	A	L	G	P	A	A	T	L	E	E	M	M	T	A	C	Q	G	V	G	G	P	G	H	K

76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
A	R	V	L	A	E	A	M	S	Q	V	T	N	S	A	T	I	M	M	Q	R	G	N	F	L

AA Distribution by Position

AMINO ACID	1	2	3	4	5	6	7	8	9
I	3	2	2	2	2	2	2	2	2
R	5	5	4	4	5	5	5	5	5
Q	6	6	6	6	6	6	6	6	6
G	6	6	6	6	5	6	6	6	6
P	5	5	5	5	5	4	4	4	3
K	6	6	6	6	6	6	5	5	5
E	7	7	7	7	7	7	7	6	6
F	2	2	2	2	2	2	2	3	3
D	3	3	3	3	3	3	3	3	3
Y	2	2	2	2	2	2	2	2	2
V	6	6	6	6	6	6	6	6	6
T	8	8	8	8	8	8	8	8	8
L	7	7	7	7	7	7	7	7	8
A	11	11	11	11	11	11	11	11	11
N	4	4	4	4	4	4	5	5	5
S	3	3	3	3	3	3	3	3	3
W	1	1	1	1	1	1	1	1	1
M	4	5	6	6	6	6	6	6	6
C	2	2	2	2	2	2	2	2	2
H	1	1	1	1	1	1	1	1	1

AA Percentage by Position

AMINO ACID	1	2	3	4	5	6	7	8	9
I	3	2	2	2	2	2	2	2	2
R	5	5	4	4	5	5	5	5	5
Q	7	7	7	7	7	7	7	7	7
G	7	7	7	7	7	7	7	7	7
P	5	5	5	5	5	4	4	4	3
K	7	7	7	7	7	7	5	5	5
E	8	8	8	8	8	8	8	7	7
F	2	2	2	2	2	2	2	3	3
D	3	3	3	3	3	3	3	3	3
Y	2	2	2	2	2	2	2	2	2
V	7	7	7	7	7	7	7	7	7
T	9	9	9	9	9	9	9	9	9
L	8	8	8	8	8	8	8	8	9
A	12	12	12	12	12	12	12	12	12
N	4	4	4	4	4	4	5	5	5
S	3	3	3	3	3	3	3	3	3
W	1	1	1	1	1	1	1	1	1
M	4	5	7	7	7	7	7	7	7
C	2	2	2	2	2	2	2	2	2
H	1	1	1	1	1	1	1	1	1

Figure 9

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
E	M	M	T	A	C	Q	G	V	G	G	P	G	H	K	A	R	V	L	A

21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
E	A	M	S	Q	V	T	N	S	A	T	I	M	M	Q	R	G	N	F	L

AA Distribution by Position

AMINO ACID	1	2	3	4	5	6	7	8	9
I	1	1	1	1	1	1	1	1	1
R	1	1	1	1	2	2	2	2	2
Q	2	2	2	3	3	3	3	2	2
G	4	4	4	4	4	5	5	5	4
P	1	1	1	1	1	1	1	1	1
K	1	1	1	1	1	1	1	1	1
E	2	1	1	1	1	1	1	1	1
V	3	3	3	3	3	3	3	3	3
T	3	3	3	3	2	2	2	2	2
L	1	1	1	1	1	1	1	1	2
A	5	5	5	5	5	4	4	4	4
N	1	1	1	1	1	1	2	2	2
S	2	2	2	2	2	2	2	2	2
M	3	4	4	3	3	3	3	3	3
C	1	1	1	1	1	1			
H	1	1	1	1	1	1	1	1	1
F								1	1

AA Percentage by Position

AMINO ACID	1	2	3	4	5	6	7	8	9
I	3	3	3	3	3	3	3	3	3
R	3	3	3	3	6	6	6	6	6
Q	6	6	6	9	9	9	9	6	6
G	13	13	13	13	13	16	16	16	13
P	3	3	3	3	3	3	3	3	3
K	3	3	3	3	3	3	3	3	3
E	6	3	3	3	3	3	3	3	3
V	9	9	9	9	9	9	9	9	9
T	9	9	9	9	6	6	6	6	6
L	3	3	3	3	3	3	3	3	6
A	16	16	16	16	16	13	13	13	13
N	3	3	3	3	3	3	6	6	6
S	6	6	6	6	6	6	6	6	6
M	9	13	13	9	9	9	9	9	9
C	3	3	3	3	3	3			
H	3	3	3	3	3	3	3	3	3
F								3	3

Figure 10

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
E	A	M	S	Q	V	T	N	S	A	T	I	M	M	Q	R	G	N	F	L

	AMINO ACID POSITION#								
	1	2	3	4	5	6	7	8	9
AMINO ACID									
I	1	1	1	1	1	1	1	1	1
R					1	1	1	1	1
Q	1	1	1	2	2	1	1	1	1
G						1	1	1	1
E	1								
V	1	1	1	1	1	1			
T	2	2	2	2	2	2	2	1	1
L									1
A	2	2	2	2	2	2	2	2	2
N	1	1	1	1	1	1	2	2	1
S	2	2	2	2	1	1	1	1	1
M	1	2	2	1	1	1	1	1	1
F								1	1

	AMINO ACID POSITION#								
	1	2	3	4	5	6	7	8	9
AMINO ACID									
I	8	8	8	8	8	8	8	8	8
R					8	8	8	8	8
Q	8	8	8	17	17	8	8	8	8
G						8	8	8	8
E	8								
V	8	8	8	8	8	8			
T	17	17	17	17	17	17	17	8	8
L									8
A	17	17	17	17	17	17	17	17	17
N	8	8	8	8	8	8	17	17	8
S	17	17	17	17	8	8	8	8	8
M	8	17	17	8	8	8	8	8	8
F								8	8

Figure 11

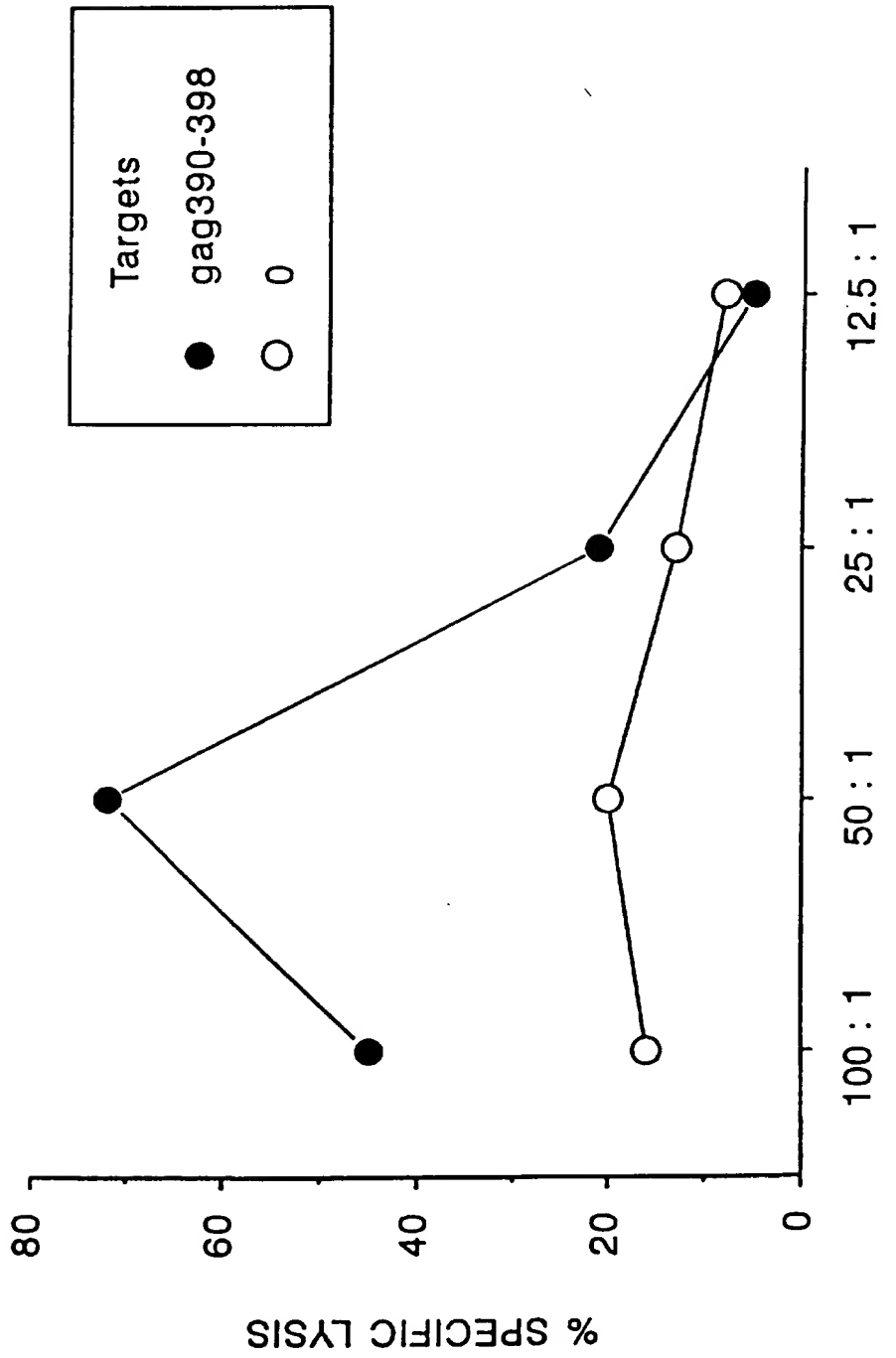
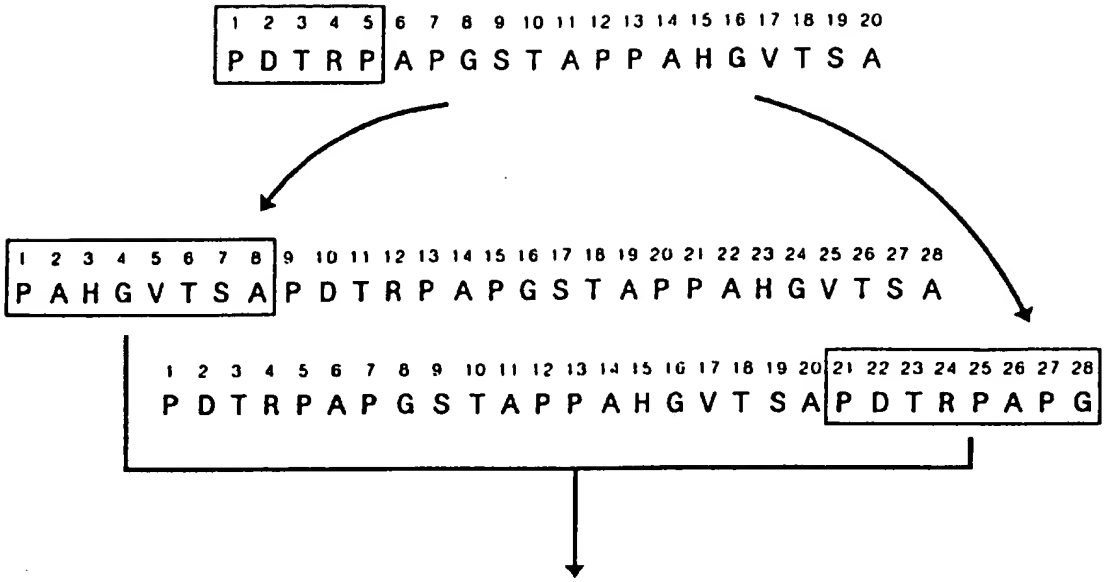

Targets
gag390-398
0
% SPECIFIC LYSIS
0
20
40
60
80
100 : 1
50 : 1
25 : 1
12.5 : 1
E:T RATIO

Figure 12



	AA Distribution by Position								
	1	2	3	4	5	6	7	8	9
AMINO ACID									
P	5	5	5	5	5	5	5	5	5
A	4	4	4	4	4	4	4	4	4
H	1	1	1	1	1	1	1	1	1
G	2	2	2	2	2	2	2	2	2
V	1	1	1	1	1	1	1	1	1
T	3	3	3	3	3	3	3	3	3
S	2	2	2	2	2	2	2	2	2
D	1	1	1	1	1	1	1	1	1
R	1	1	1	1	1	1	1	1	1

	AA Percentage by Position								
	1	2	3	4	5	6	7	8	9
AMINO ACID									
P	25	25	25	25	25	25	25	25	25
A	20	20	20	20	20	20	20	20	20
H	5	5	5	5	5	5	5	5	5
G	10	10	10	10	10	10	10	10	10
V	5	5	5	5	5	5	5	5	5
T	15	15	15	15	15	15	15	15	15
S	10	10	10	10	10	10	10	10	10
D	5	5	5	5	5	5	5	5	5
R	5	5	5	5	5	5	5	5	5

Figure 13

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
C	T	Y	S	P	A	L	N	K	M	F	C	Q	L	A	K	T	C	P	V	Q	L	W	V	D	S	T	P
								M	L	L	Y																

AA Distribution by Position

	1	2	3	4	5	6	7	8	9	10
AMINO ACID										
C	3	2	2	2	2	2	2	2	2	2
T	2	2	1	1	1	1	1	1	2	2
Y	2	2	2	1	1	1	1	1	1	1
S	1	1	1	1				1	1	1
P	2	2	2	2	2	1	1	1	1	2
A	2	2	2	2	2	2	1	1	1	1
L	4	4	4	5	5	5	5	4	4	4
N	1	1	1	1	1	1	1	1		
K	2	2	2	2	2	2	2	2	2	1
M	2	2	2	2	2	2	2	2	2	1
F	1	1	1	1	1	1	1	1	1	1
Q	1	1	2	2	2	2	2	2	2	2
V		1	1	1	1	2	2	2	2	2
W					1	1	1	1	1	1
D							1	1	1	1

AA Percentage by Position

	1	2	3	4	5	6	7	8	9	10
AMINO ACID										
C	13	9	9	9	9	9	9	9	9	11
T	9	9	4	4	4	4	4	4	9	11
Y	9	9	9	4	4	4	4	4	4	5
S	4	4	4	4				4	4	5
P	9	9	9	9	9	4	4	4	4	11
A	9	9	9	9	9	9	4	4	4	5
L	17	17	17	22	22	22	22	17	17	18
N	4	4	4	4	4	4	4	4		
K	9	9	9	9	9	9	9	9	9	5
M	9	9	9	9	9	9	9	9	9	5
F	4	4	4	4	4	4	4	4	4	5
Q	4	4	9	9	9	9	9	9	9	11
V		4	4	4	4	9	9	9	9	11
W					4	4	4	4	4	5
D							4	4	4	5

Figure 14

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
W	H	S	N	L	N	D	A	T	Y	Q	R	T	R	A	L	V	R	T	G	M	D	P	R	M

AA Distribution by Position

	1	2	3	4	5	6	7	8	9
AMINO ACID									
W	1								
H	1	1							
S	1	1	1						
N	2	2	2	2	1	1			
L	2	2	2	2	2	1	1	1	1
D	1	1	1	1	1	1	2	1	1
A	2	2	2	2	2	2	2	2	1
T	2	2	2	3	3	3	3	3	3
Y	1	1	1	1	1	1	1	1	1
Q	1	1	1	1	1	1	1	1	1
R	2	2	3	3	3	3	3	3	4
V		1	1	1	1	1	1	1	1
G					1	1	1	1	1
M						1	1	1	1
P								1	1

AA Percentage by Position

	1	2	3	4	5	6	7	8	9
AMINO ACID									
W	6								
H	6	6							
S	6	6	6						
N	13	13	13	13	6	6			
L	13	13	13	13	13	6	6	6	7
D	6	6	6	6	6	6	13	6	7
A	13	13	13	13	13	13	13	13	7
T	13	13	13	19	19	19	19	19	13
Y	6	6	6	6	6	6	6	6	7
Q	6	6	6	6	6	6	6	6	7
R	13	13	19	19	19	19	19	19	25
V		6	6	6	6	6	6	6	7
G					6	6	6	6	7
M						6	6	6	7
P								6	7

Figure 15

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
S	D	Y	E	G	R	L	I	T	Y	Q	R	T	R	A	L	V	A	S	N	E	N	M	E	T	M

AA Distribution by Position

	1	2	3	4	5	6	7	8	9
AMINO ACID									
S	1	1	1	1	1	1	1	1	1
D	2	2	1	1	1	1	1	1	1
Y	2	2	2	1	1	1	1	1	1
E	1	1	1	2	1	1	2	2	2
G	1	1	1	1	1				
R	3	3	3	3	3	3	2	2	2
L	2	2	2	2	2	2	2	1	1
I	1	1	1	1	1	1	1	1	
T	2	2	2	2	2	2	2	3	3
Q	1	1	1	1	1	1	1	1	1
A	2	2	2	2	2	2	2	2	2
V	1	1	1	1	1	1	1	1	1
N			1	1	2	2	2	2	2
M						1	1	1	2

AA Percentage by Position

	1	2	3	4	5	6	7	8	9
AMINO ACID									
S	5	5	5	5	5	5	5	5	5
D	11	11	5	5	5	5	5	5	5
Y	11	11	11	5	5	5	5	5	5
E	5	5	5	11	5	5	11	11	11
G	5	5	5	5	5				
R	16	16	16	16	16	16	11	11	11
L	11	11	11	11	11	11	11	5	5
I	5	5	5	5	5	5	5	5	
T	11	11	11	11	11	11	11	16	16
Q	5	5	5	5	5	5	5	5	5
A	11	11	11	11	11	11	11	11	11
V	5	5	5	5	5	5	5	5	5
N			5	5	11	11	11	11	11
M						5	5	5	11

Figure 16